Quantum algorithms for hidden nonlinear structures

Andrew M. Childs
amchilds@caltech.edu

Leonard J. Schulman
schulman@caltech.edu

Umesh V. Vazirani
vazirani@cs.berkeley.edu

Abstract

Attempts to find new quantum algorithms that outperform classical computation have focused primarily on the nonabelian hidden subgroup problem, which generalizes the central problem solved by Shor's factoring algorithm. We suggest an alternative generalization, namely to problems of finding hidden nonlinear structures over finite fields. We give examples of two such problems that can be solved efficiently by a quantum computer, but not by a classical computer. We also give some positive results on the quantum query complexity of finding hidden nonlinear structures.

1 Introduction

One of the major open problems in quantum computation is to develop new quantum algorithms. Much of the work on this question has focused on the nonabelian hidden subgroup problem (HSP), attempting to extend the quantum solution of the abelian HSP [11, 17, 18]. Unfortunately, these efforts have met with only limited success. In this paper, we describe an alternative way of generalizing the success of Shor's algorithm.

The key to exponential savings in quantum algorithms is the creation of sharp constructive interference in large sets. Such precise interference is only known to arise in a few cases, primarily in which the set is a group. Under these conditions, the key to quantum speed-up is to diagonalize the group algebra, i.e., to perform a Fourier transform. Once this has been done, certain structures become easy to detect.

The structures that have been investigated so far are subgroups and their cosets. In the case of abelian groups, the Fourier transform is a mapping from the group to its dual, and this mapping respects subgroups and cosets. Advances in quantum algorithms have been pursued by extending the groups from abelian to nonabelian, but in the nonabelian case there is no dual group, and the same approach is not available. Indeed, certain methods that work in the abelian case are known to fail in some nonabelian cases, such as the symmetric group [7, 8, 13].

Our approach in this paper is to shift the focus back to the Fourier transform over abelian groups, and to consider what other hidden structures can be revealed by abelian Fourier transforms via constructive interference effects. We turn for inspiration to optics and acoustics, where light or sound can be highly focused (i.e., undergo highly constructive interference) when reflected by a conic (e.g., parabolic or elliptic) surface. To connect this idea to known quantum algorithms, observe that abelian hidden subgroup problems, when restricted to a vector space, can be viewed as determining a hidden linear structure. (Most generally, this viewpoint makes sense for a module over any ring, but here we restrict ourselves to vector spaces over finite fields.) Any subgroup of the additive group of $\mathbb{F}_q^d$ ($q = p^m$ a prime power) is an $\mathbb{F}_q$-linear subspace, and the cosets of this subgroup consist of parallel affine subspaces, or *flats*. Given a black box function that is constant on each flat and distinct on different ones, abelian Fourier sampling determines the hidden subspace in time $\mathrm{poly}(d \log q)$. Pursuing the analogy with wave mechanics, our approach is to set up black box functions that are constant on quadratic surfaces, and use interference effects to discover properties of the unknown quadratic. More generally, we will study this approach for algebraic sets of higher degree.

The first problem we study is the *hidden radius problem*. In this problem, the hidden property is the radius r of a sphere. We give an efficient quantum algorithm for determining one bit of r, namely whether or not it is a quadratic residue, assuming that the dimension is odd. With a classical computation, even this restricted problem requires exponentially many queries. (For the problem of determining the other bits of r, we argue that the quantum query complexity is small.)

The second problem we discuss is the *hidden flat of centers problem*. In this problem, the radius of the sphere is fixed (say, at $r = 1$), but its center is constrained to lie in an unknown flat in $\mathbb{F}_q^d$. For example, the centers of the spheres may lie on an unknown line. For this problem, we give an efficient quantum algorithm to determine the entire hidden flat, not just one bit of information about it. However, this algorithm also works only when the dimension is odd. The main idea of the algorithm is to use a quantum walk to move amplitude from the spheres to their centers. Our algorithms for both this and the hidden radius problem make crucial use of a connection to certain exponential sums called *twisted Kloosterman sums*.

Both of the above problems fall into a framework of *shifted subset problems*. For problems in this class, the main idea is to define a black box function that is constant on some subset of the points in $\mathbb{F}_q^d$, as well as on shifted versions of this subset, with the function taking distinct values when the shifts are different. The goal may be either to determine some property of the basic subset, or of the allowed shifts, or both. Typically, this will not give a well-defined black box, since different shifts of the subset may lead to overlapping points. However, we can resolve this issue by defining the black box carefully.

We also obtain results regarding hidden polynomial structures of higher degree. These results are purely information-theoretic (i.e., regard query complexity). We introduce the framework of *hidden polynomial problems*. In these problems, the hidden object is a multivariate polynomial $h(x) \in \mathbb{F}_q[x_1, \ldots, x_d]$ chosen from some set of possible polynomials. We are given a black box function that is constant on the level sets of $h(x)$ (i.e., the sets $\{x \in \mathbb{F}_q^d : h(x) = y\}$ for various $y \in \mathbb{F}_q$) and distinct on different level sets, and the goal is to determine $h(x)$. When $h(x)$ is linear, this is the abelian HSP described above, whereas for more general polynomials, it is typically not an HSP in any group. (Observe that a hidden polynomial problem, unlike a shifted subset problem, is automatically an oracle problem.)

Assuming the dimension d and the degree of $h(x)$ is constant, we show that the query complexity of the hidden polynomial problem is typically $\mathrm{poly}(\log q)$. We show this by considering an analog of the standard approach to the HSP, wherein one query of the black box is used to produce a quantum state that depends on the hidden object. Provided these states are sufficiently statistically distinguishable, it follows that $\mathrm{poly}(\log q)$ copies contain enough information to determine the hidden object with high probability. To establish distinguishability of the states, we give two simple but apparently new results about the fidelity between general quantum states satisfying certain intersection conditions, and we show that one of these conditions is satisfied by typical polynomials. These lemmas could also have applications to problems involving quantum states derived from combinatorial designs that are unrelated to polynomials.

2 Hidden radius problem

We begin by considering the first of two shifted subset problems, the *hidden radius problem.* In the quantum version of the hidden radius problem, our goal is to determine an unknown radius $r \in \mathbb{F}_q$ given a uniform superposition over points in $\mathbb{F}_q^d$ on a sphere of radius r whose center is chosen uniformly at random. We give an efficient quantum algorithm for determining whether r is a quadratic residue, provided d is odd. We also show that the quantum query complexity of finding r is $\operatorname{poly}(\log q)$, again assuming d is odd, and we give evidence that this should also be the case for d even.

For this problem to make sense classically as well as quantumly, we define it in terms of a black box function. Roughly speaking, we would like to define a black box function that on input x, a point on the sphere of radius r with center t, outputs some encryption of t, thereby giving a function that is constant on shifted spheres and distinct on different spheres. But this cannot be done directly, since spheres can intersect. To resolve this issue, we note that the vector $s = x - t$ pointing to x from the center t uniquely describes a particular sphere. So our black box f_1 takes as input the pair x and an encryption σ of s and outputs an encryption of t. We also supply a black box f_{-1} that takes as input the pair x and encryption of t and outputs the encryption of s. The goal of the problem is to determine r using an oracle that computes either f_1 or f_{-1} as desired. (In Appendix A, we give a black-box formulation of general shifted subset problems, which provides an alternative oracle for the hidden radius problem.)

It is straightforward to show that this problem is hard for a classical computer.

Theorem 1. *Any classical computation with access to f_1 and f_{-1} requires an expected exponential number of queries to obtain a $1/\operatorname{poly}(d \log q)$ bias for any single bit of information about r.*

Proof sketch. Let the hidden radius r be uniformly random, and let $f_1(x, \sigma)$ be a uniformly random one-to-one function of the sphere center $t = x - s$, where σ is the encryption of s. Now we can assume without loss of generality that the algorithm is deterministic. Given any sequence of evaluations of f_1 and f_{-1} that do not involve any sphere twice, the conditional distribution on subsequent evaluations involving points on other spheres is uniform. The probability of success is therefore sub-polynomial so long as the square of the number of queries is less than a polynomial fraction of q^d. □

To solve this problem on a quantum computer, we use the following state generation procedure. Begin with a uniform superposition over x and σ, then compute f_1, then uncompute σ using f_{-1}, and finally discard the function value, giving (up to normalization)

$$\begin{aligned}
\sum_{x,\sigma} |x, \sigma\rangle &\mapsto \sum_{x,\sigma} |x, \sigma, f_1(x, \sigma)\rangle && (1) \\
&\mapsto \sum_{x,\sigma} |x, f_1(x, \sigma)\rangle && (2) \\
&\mapsto |\mathcal{S}_r + t\rangle \text{ where } t \text{ is uniformly random in } \mathbb{F}_q^d && (3)
\end{aligned}$$

where $\mathcal{S}_r$ denotes the sphere of radius r centered at the origin, and where we use the convention that for a finite set S, $|S\rangle := \sum_{s \in S} |s\rangle / \sqrt{|S|}$ denotes the normalized uniform superposition over elements of S. In other words, we can use two queries to the oracle to produce the mixed quantum state

$$\rho_r := \frac{1}{q^d} \sum_{t \in \mathbb{F}_q} |\mathcal{S}_r + t\rangle\langle\mathcal{S}_r + t| \qquad (4)$$

from which we would like to extract information about the hidden radius r.

The sphere of radius r centered at the origin is defined by $\mathcal{S}_r := L_{\Delta(x),r}$, where $\Delta(x) := \sum_{j=1}^{d} x_j^2$, and where $L_{f,y} := f^{-1}(y) = \{x \in X : f(x) = y\}$ denotes the level set of $f(x)$ with value y. The quadratic polynomial $\Delta(x)$ can be thought of as measuring the distance from the origin in $\mathbb{F}_q^d$ (although of course it does not satisfy a triangle inequality); then $\mathcal{S}_r$ consists of the points at distance r from the origin.

Note that since we are working in a finite field, there is no concept of large spheres or small spheres; indeed all spheres contain approximately the same number of points. In particular, the number of points on the sphere of radius r is [12, Theorem 1]

$$|\mathcal{S}_r| = \begin{cases} q^{d-1} + \chi((-1)^{(d-1)/2} r)\sqrt{q^{d-1}} & d \text{ odd}, r \neq 0 \\ q^{d-1} - \chi((-1)^{d/2})\sqrt{q^{d-2}} & d \text{ even}, r \neq 0 \\ q^{d-1} & d \text{ odd}, r = 0 \\ q^{d-1} + \chi((-1)^{d/2})(q-1)\sqrt{q^{d-2}} & d \text{ even}, r = 0\,, \end{cases} \tag{5}$$

where χ denotes the quadratic character of $\mathbb{F}_q^\times$. In other words, up to small corrections, every sphere has about q^{d-1} points on it (except that the sphere of zero radius in two dimensions consists of $2q-1$ points when $q = 1 \bmod 4$; and is simply a single point, the origin, when $q = 3 \bmod 4$).

Our goal is to determine r using polynomially many copies of the hidden radius state ρ_r. For any r, the state is invariant under arbitrary translations in $\mathbb{F}_q^d$. This symmetry can be exploited using the d-dimensional Fourier transform over $\mathbb{F}_q$,

$$U := \frac{1}{\sqrt{q^d}} \sum_{x,k \in \mathbb{F}_q^d} \omega_p^{\operatorname{tr} k\cdot x} |k\rangle\langle x|\,, \tag{6}$$

where $\omega_p := e^{2\pi i/p}$, $k \cdot x := \sum_{j=1}^{d} k_j x_j$, and where $\operatorname{tr} a := a + a^p + \cdots + a^{q/p}$ denotes the trace from $\mathbb{F}_q$ to $\mathbb{F}_p$. Fourier transforming the state, we find

$$U \rho_r U^\dagger = \sum_{k \in \mathbb{F}_q^d} \Pr(k|r)\, |k\rangle\langle k| \qquad \text{with} \qquad \Pr(k|r) = \frac{1}{q^d |\mathcal{S}_r|} \left| \sum_{x \in \mathcal{S}_r} \omega_p^{\operatorname{tr} k\cdot x} \right|^2 . \tag{7}$$

Since the resulting density matrix is diagonal, we can measure in the Fourier basis without loss of information, and all that remains is to infer r from samples of $\Pr(k|r)$.

To understand this distribution, we must understand the Fourier transform of a sphere, which is given by [12]

$$\sum_{x \in \mathcal{S}_r} \omega_p^{\operatorname{tr} k\cdot x} = \frac{G_1^d}{q} K_{\chi^d}(r, \Delta(k)/4) \tag{8}$$

(assuming $k \neq 0$), where $G_1 = -(-1)^m \sqrt{q}$ when $p = 1 \bmod 4$, and $G_1 = -(-i)^m \sqrt{q}$ when $p = 3 \bmod 4$, and where we define the *η-twisted Kloosterman sum*

$$K_\eta(a,b) := \sum_{c \in \mathbb{F}_q} \eta(c)\, \omega_p^{\operatorname{tr}(ac + bc^{-1})} \tag{9}$$

for $a, b \in \mathbb{F}_q$, and for any multiplicative character η of $\mathbb{F}_q^\times$. (This exponential sum can be viewed as the discrete analog of a Bessel function.)

If the dimension is odd, then we are interested in a χ-twisted Kloosterman sum, also known as a Salié sum. This has the explicit form [4, 15]

$$K_\chi(a,b) = \begin{cases} G_1 & ab = 0,\ a \neq 0 \text{ or } b \neq 0 \\ 2\chi(b) G_1 \cos \frac{4\pi \operatorname{tr} \sqrt{ab}}{p} & \chi(ab) = 1 \\ 0 & \chi(ab) = -1 \text{ or } a = b = 0\,. \end{cases} \tag{10}$$

In particular, we see that $\Pr(\Delta(k) = 0)$ is exponentially small, and for $\Delta(k) \neq 0$, $\chi(\Delta(k))$ determines $\chi(r)$ as follows. For $r \neq 0$, if $\chi(r\Delta(k)) = -1$, $\Pr(k|r) = 0$. On the other hand, if $r = 0$, $\Pr(\chi(\Delta(k)) = +1) = \Pr(\chi(\Delta(k)) = -1) = 1/2 - o(1)$. This gives a simple quantum algorithm to determine $\chi(r)$.

Theorem 2. *For d odd, there is an efficient bounded-error quantum algorithm to determine $\chi(r)$.*

Proof. The algorithm repeats the following process a constant number of times: Prepare ρ_r, perform the Fourier transform, measure a value of k, compute $\Delta(k)$, and discard the result if $\Delta(k) = 0$. If the results include points with both $\chi(\Delta(k)) = +1$ and $\chi(\Delta(k)) = -1$, output $r = 0$. Otherwise, output the common value of $\chi(\Delta(k))$. A straightforward calculation shows that this algorithm succeeds with constant probability. □

Ideally, we would like to determine not just $\chi(r)$, but rather r itself. While we do not know an efficient algorithm, we can at least show that polynomially many queries suffice:

Theorem 3. *For d odd,* $\operatorname{poly}(\log q)$ *queries to the hidden radius oracle suffice to determine r.*

The proof is given in Appendix B.

If the dimension is even, then the distribution $\Pr(k|r)$ depends on the (non-twisted) Kloosterman sum

$$K_1(a,b) = \sum_{c \in \mathbb{F}_q} \omega_p^{\operatorname{tr}(ac + bc^{-1})} = \sum_{c \in \mathbb{F}_q} \chi(c^2 - 4ab)\, \omega_p^c\,. \tag{11}$$

No closed-form expression for such sums is known. But we do know that in the limit $q \to \infty$, the distribution of values of the Kloosterman sum asymptotically approaches the Sato-Tate (semicircle) distribution [1, 10], and indeed the convergence to this distribution is rapid [14]. Since the Sato-Tate distribution is far from uniform, this shows that the states $\rho_r, \rho_{r'}$ are information-theoretically distinguishable for typical pairs $r \neq r'$. We conjecture that in fact arbitrary pairs can be distinguished.

Not only do we not have a closed-form expression for non-twisted Kloosterman sums, but we do not even know whether they can be efficiently approximated on a quantum computer. If we could approximate these sums, then we could efficiently distinguish distinguishable pairs of radii. The problem of approximately computing Kloosterman sums (as well as more general exponential sums) on a quantum computer appears to be a natural open problem. Indeed, it will also be relevant to the even-dimension case of the problem considered in the following section.

3 Hidden flat of centers problem

In this section, we consider a second shifted subset problem, the *hidden flat of centers problem.* In this problem, unlike the hidden radius problem, the spheres are promised to have unit radius. Their centers lie on an unknown flat H, and the goal is to determine this flat. For a general black-box formulation of shifted subset problems that applies to the hidden flat of centers problem, see

Appendix A. With that black box, the classical query complexity of determining H is exponential in $d \log q$. Here we give an efficient quantum algorithm for finding H, provided $d = O(1)$ is odd.

Using the quantum oracle for the hidden flat of centers problem, we can produce the quantum state

$$\rho_H := \frac{1}{|H|} \sum_{h \in H} |\mathcal{S}_1 + h\rangle\langle \mathcal{S}_1 + h| \,. \tag{12}$$

Our goal is to determine H by making measurements on this state. We do this by using a quantum walk to move amplitude from $\mathcal{S}_1 + h$ to h. If we can move a sufficiently large fraction of the amplitude, then we can determine the hidden flat by (classically) solving a noisy linear algebra problem.

To move amplitude from unit spheres to their centers, we will use a continuous-time quantum walk on the *Winnie Li graph.* This graph has vertex set $\mathbb{F}_q^d$, and edges between points $x, x' \in \mathbb{F}_q^d$ with $\Delta(x - x') = 1$. Thus its adjacency matrix is

$$A := \sum_{x \in \mathbb{F}_q^d} \sum_{s \in \mathcal{S}_1} |x + s\rangle\langle x| \,. \tag{13}$$

The continuous-time quantum walk for time t is simply the unitary operator e^{-iAt}. This unitary operator can be efficiently implemented on a quantum computer provided we can efficiently transform into the eigenbasis of A, and can efficiently compute the eigenvalue corresponding to a given eigenvector.

The adjacency matrix (13) has eigenvectors

$$|\tilde{k}\rangle := \frac{1}{\sqrt{q^d}} \sum_{x \in \mathbb{F}_q^d} \omega_p^{\operatorname{tr} k \cdot x} |x\rangle \tag{14}$$

for $k \in \mathbb{F}_q^d$, as is clear from translation invariance. Thus we can transform to the eigenbasis of A simply using the Fourier transform (6). The corresponding eigenvalues are given by the Fourier transform of a unit sphere (cf. Section 2):

$$\lambda_k = \sum_{x \in \mathcal{S}_1} \omega_p^{\operatorname{tr} k \cdot x} = \begin{cases} |\mathcal{S}_1| & k = 0 \\ G_1^d K(1, \Delta(k)/4)/q & \text{otherwise} \,. \end{cases} \tag{15}$$

All of these eigenvalues are $O(\sqrt{q^{d-1}})$, with the exception of $\lambda_0 = \Theta(q^{d-1})$. It will be helpful to remove the single large eigenvalue, so we will replace A by $\bar{A} := A - \lambda_0 |\tilde{0}\rangle\langle\tilde{0}|$. Then we have $\|\bar{A}\| \le 2\sqrt{q^{d-1}}$ [19].

Lemma 4. *Suppose we start with the quantum state (12), perform the quantum walk with the modified adjacency matrix $\bar{A}$ for time $t = 1/\sqrt{q^{d-1}} \log q$, and finally measure in the computational basis. Then each point in H occurs with probability $|H|^{-1}[1/\log q + O(1/\log^{3/2} q)]$, and any point not on H occurs with probability $O(q^{-d})$.*

Proof. Consider the evolution of a single sphere $|\mathcal{S}_1 + x\rangle$. Taylor expanding the action of the walk, the amplitude at the center x is

$$\begin{aligned} \langle x|e^{-i\bar{A}t}|\mathcal{S}_1 + x\rangle &= -it\langle x|\bar{A}|\mathcal{S}_1 + x\rangle + O(\|\bar{A}\|^2 t^2) \qquad (16) \\ &= -it\sqrt{|\mathcal{S}_1|}[1 - O(q^{-1})] + O(\|\bar{A}\|^2 t^2) \,, \qquad (17) \end{aligned}$$

so

$$|\langle x|e^{-i\bar{A}t}|\mathcal{S}_1+x\rangle|^2 = \frac{1}{\log q} + O(\log^{-3/2} q)\,. \tag{18}$$

Averaging over $x \in H$ gives the first part of the lemma.

It remains to show that the background is nearly uniform. To see this, note that $e^{-i\bar{A}t}$ leaves invariant the subspace $\mathrm{span}\{|x\rangle, |\mathcal{S}_0+x\rangle, |\mathcal{S}_1+x\rangle, \ldots, |\mathcal{S}_{q-1}+x\rangle\}$ (which contains the state $|\tilde{0}\rangle$), since $A|x\rangle = \sqrt{|\mathcal{S}_1|}|\mathcal{S}_1+x\rangle$ and

$$A|\mathcal{S}_r+x\rangle = \frac{1}{\sqrt{|\mathcal{S}_r|}}\sum_{s\in\mathcal{S}_r}\sum_{s'\in\mathcal{S}_1}|s+s'+x\rangle = \frac{1}{\sqrt{|\mathcal{S}_r|}}\sum_{y\in\mathbb{F}_q^d}|\mathcal{S}_1\cap\mathcal{S}_r+y-x|\,|y\rangle\,. \tag{19}$$

Here the coefficient of y depends only on $\Delta(y-x)$ (with $y=x$ a special case distinct from $\Delta(y-x)=0$) by the fact that the orthogonal group over $\mathbb{F}_q^d$ acts transitively on nonzero points of fixed norm (as a consequence of Witt's Lemma [2]). Thus the evolved state $e^{-i\bar{A}t}|\mathcal{S}_1+x\rangle$ is spherically symmetric about x.

Now consider making a measurement on $|\mathcal{S}_r+x\rangle$ for some $r\in\mathbb{F}_q$: each point on $\mathcal{S}_r+x$ occurs with probability $1/|\mathcal{S}_r|$. Averaging over $x\in H$, we see that the point $y\in\mathbb{F}_q^d$ occurs with probability $|\{x\in H : y\in\mathcal{S}_r+x\}|/(|\mathcal{S}_r|\cdot|H|)$. Since $|\mathcal{S}_r|=\Theta(q^{d-1})$, $|H|=q^{\dim H}$, and the numerator is $|H\cap(\mathcal{S}_1+y)|=O(q^{\dim H-1})$, the probability of seeing any $y\in\mathbb{F}_q^d$ is $O(q^{-d})$. Thus the piece of $e^{-i\bar{A}t}|\mathcal{S}_1+x\rangle$ orthogonal to $|x\rangle$, when averaged over $x\in H$, contributes probability $O(q^{-d})$ to every point $y\in\mathbb{F}_q^d$. □

Now we show how to reconstruct the flat H using samples from this distribution. A priori, $\dim H$ is unknown, so we iteratively try increasing values of $\dim H$ until the following procedure identifies H.

Let $d'=\dim H+1$, so that d' points in affine general position are sufficient to determine H. Suppose we sample $k=\mathrm{poly}(\log q)$ points, so that with high probability the number of points in H is at least $4d'$. The following lemma shows that with high probability the k-sample does not intersect any flat H' other than H, of the same dimension as H, in more than $4d'$ points. Thus the flat H can be computed by exhaustively trying all $\binom{k}{4d'}=\mathrm{poly}(\log q)$ subsets of the sample points.

Lemma 5. *Suppose we sample k points independently and identically with the following distribution: the point is uniformly random in H with probability at least $1/\mathrm{poly}(\log q)$, and any point not in H has probability at most c/q^d for some constant c. Then $\Pr[\exists H'\neq H, \dim H'=\dim H$, with $\geq 4d'$ points from the k-sample$]\leq O(\binom{k}{d'}^2)(c/q)^{d'}$.*

Proof sketch. For this event to occur, either $2d'$ points must fall in $H'\cap H$ or $2d'$ points must fall in $H'-H$. We bound the probabilities of each of these events by similar arguments.

Consider the first of these events. Let $s_1,\ldots,s_{2d'}$ be the first $2d'$ points of the k-sample that fall in $H'\cap H$. Since $\Pr[\dim\mathrm{affspan}\{s_1,\ldots,s_{2d'}\}\leq d'-2]\leq(1+O(1/q))\Pr[\dim\mathrm{affspan}\{s_1,\ldots,s_{2d'}\}\leq d'-2$ and $\dim\mathrm{affspan}\{s_1,\ldots,s_{d'}\}=\dim\mathrm{affspan}\{s_1,\ldots,s_{2d'}\}]$ (where affspan denotes the affine span of a set of points), it is sufficient to bound the probability of the latter event. For each of the $\binom{k}{d'}$ subsets $s_1,\ldots,s_{d'}$ within the k-sample, the probability of this event is bounded by the number of ways of choosing the remaining d' points out of k, times the probability that all remaining d' points fall in $\mathrm{affspan}\{s_1,\ldots,s_{d'}\}$. Overall the probability of this event is bounded by $\binom{k}{d'}^2(1/q)^{d'}$.

The case of $H'-H$ is similar; the only change is that because we have less control over the probabilities with which points are selected, the final bound is $\binom{k}{d'}^2(c/q)^{d'}$. □

Overall, we have shown

Theorem 6. *Suppose $d = O(1)$ is odd. Then there is a quantum algorithm to determine the hidden flat of centers in time* $\operatorname{poly}(\log q)$.

Note that we could use the same algorithm for d even, provided we could efficiently approximate the eigenvalues of A by approximately calculating (non-twisted) Kloosterman sums.

4 Hidden polynomial problems

In this section, we prove some general results on the distinguishability of black box functions, and then use these results to show that the quantum query complexity of the hidden polynomial problem (defined in Section 4.2) is typically polynomial.

4.1 Distinguishability of states with given intersection properties

Consider a black box function $f : X \to Y$ where X, Y are finite sets. Any such function can be encoded in a quantum state using an approach analogous to the so-called standard method for the hidden subgroup problem. In this approach, we begin with a uniform superposition over the input space, compute the black box function in an auxiliary register, and then discard that register, giving (up to normalization)

$$\sum_{x \in X} |x\rangle \mapsto \sum_{x \in X} \sum_{x \in X} |x, f(x)\rangle \tag{20}$$

$$\mapsto |L_{f,y}\rangle \text{ where } y \in Y \text{ occurs with probability } |L_{f,y}|/|X| \,. \tag{21}$$

(Recall that $L_{f,y} := f^{-1}(y) = \{x \in X : f(x) = y\}$ denotes the level set of $f(x)$ with value y.) In other words, this procedure uses one query of the black box to produce the mixed quantum state

$$\rho_f := \sum_{y \in Y} \frac{|L_{f,y}|}{|X|} |L_{f,y}\rangle\langle L_{f,y}| \,. \tag{22}$$

Suppose that f is chosen from a set $\mathcal{F}$ of possible black box functions (where $\log |\mathcal{F}| = \operatorname{poly}(\log |X|)$), and we would like to determine which one we have. Then we can create $t = \operatorname{poly}(\log |X|)$ copies of the state (22), $\rho_f^{\otimes t}$, and perform a quantum measurement to attempt to determine f. If some such measurement succeeds with high probability, then the query complexity of the problem is polynomial. For some measurement to succeed, it suffices to show that the single-copy states are pairwise distinguishable, as measured by the quantum fidelity

$$F(\rho, \rho') := \operatorname{tr} |\sqrt{\rho}\sqrt{\rho'}| \,. \tag{23}$$

This follows from a result of Barnum and Knill [3]:

Theorem 7. *Suppose ρ is drawn from an ensemble $\{\rho_1, \ldots, \rho_N\}$, where each ρ_k occurs with some fixed prior probability. Then there exists a quantum measurement that returns the outcome k with probability at least $1 - N\sqrt{\max_{i \neq j} F(\rho_i, \rho_j)}$.*

(In fact, by the minimax theorem, this result holds even without assuming a prior distribution for the ensemble [9].) In particular, since $F(\rho^{\otimes \ell}, \rho'^{\otimes \ell}) = F(\rho, \rho')^{\ell}$, arbitrarily small error probability $\epsilon > 0$ can be achieved using $\ell \geq \lceil 2(\log N - \log \epsilon)/\log(1/\max_{i \neq j} F(\rho_i, \rho_j)) \rceil$, so $\ell = \operatorname{poly}(\log N)$ copies suffice provided the maximum fidelity is bounded away from 1 by at least $1/\operatorname{poly}(\log N)$.

Such an argument has been used to show that the query complexity of the hidden subgroup problem is polynomial [5]; here we give analogous results for the hidden polynomial problem.

We begin by giving two bounds on the pairwise fidelity in terms of the intersection properties of the functions.

Lemma 8. *Suppose* $\Pr_{y,y'\in Y}[|L_{f,y} \cap L_{f',y'}| \geq \alpha] \leq \beta$*, and* $|L_{f,y}| \leq \delta$ *for all* $y \in Y$*. Then* $F(\rho_f, \rho_{f'})^2 \leq (\alpha^2 + \beta\delta^2)|Y|^3/|X|^2$*.*

Proof. By the Cauchy-Schwartz inequality applied to the singular values of $\sqrt{\rho_f}\sqrt{\rho_{f'}}$ (whose rank is clearly at most $|Y|$),

$$F(\rho_f, \rho_{f'})^2 \leq |Y| \operatorname{tr} \rho_f \rho_{f'} \tag{24}$$

$$= \frac{|Y|}{|X|^2} \sum_{y,y'\in Y} |L_{f,y} \cap L_{f',y'}|^2 \,, \tag{25}$$

and the claim follows from the assumptions. □

Lemma 9. *Suppose* $\Pr_{y\in Y}[|L_{f,y} \cap L_{f',y'}| \geq \alpha] \leq \beta$ *for all* $y' \in Y$*, and* $\gamma \leq |L_{f,y}| \leq \delta$ *for all* y*. Then* $F(\rho_f, \rho_{f'})^2 \leq \alpha|Y|^2/\gamma|X| + \beta\delta|Y|/|X|$*.*

Proof. Let Π_ρ denote the projector onto the support of ρ. By considering the POVM with elements $\Pi_\rho, 1 - \Pi_\rho$ and noting that the classical fidelity of the resulting distribution is an upper bound on the quantum fidelity, we have $F(\rho, \rho') \leq \sqrt{\operatorname{tr} \Pi_\rho \rho'}$. Thus

$$F(\rho_f, \rho_{f'})^2 \leq \frac{1}{|X|} \sum_{y,y'\in Y} \frac{|L_{f,y} \cap L_{f',y'}|^2}{|L_{f,y}|} \tag{26}$$

$$\leq \frac{\alpha|Y|^2}{\gamma|X|} + \frac{1}{|X|} \sum_{y\in Y_{\text{bad}}} \frac{\left(\sum_{y'\in Y} |L_{f,y} \cap L_{f',y'}|\right)^2}{|L_{f,y}|} \tag{27}$$

$$\leq \frac{\alpha|Y|^2}{\gamma|X|} + \frac{1}{|X|} \sum_{y\in Y_{\text{bad}}} |L_{f,y}| \tag{28}$$

where $Y_{\text{bad}} := \{y \in Y : |L_{f,y} \cap L_{f',y'}| \geq \alpha$ for some $y' \in Y\}$. Then the claim follows from the assumptions. □

4.2 Distinguishability of hidden polynomial states

Now we specialize to the hidden polynomial problem. Let $h(x) \in \mathbb{F}_q[x_1, \ldots, x_d]$ be a polynomial in d variables over $\mathbb{F}_q$ of total degree $\deg h = O(1)$. This polynomial is hidden by a function $f : X \to Y$ where $X = \mathbb{F}_q^d$ and $|Y| \geq q$, which is simply h composed with an arbitrary injective function from $\mathbb{F}_q$ to Y. In particular, the level sets of f are isomorphic to the level sets of h. It is important that the black box hiding function f is not simply the hidden polynomial h, so that the problem of reconstructing h from queries to f will be hard for a classical computer. However, $\rho_f = \rho_h$ by the isomorphism of the level sets, so it is sufficient to calculate the fidelity between the states as if the hiding functions were in fact the polynomials.

We begin by specializing Lemmas 8 and 9 to the case of hidden polynomial states. Here and in what follows, the number of variables and the degrees of polynomials are considered bounded; the notation $o(1)$ is with respect to the limit $q \to \infty$.

Corollary 10. *Let $d \geq 2$, and suppose* $\Pr_{y,y' \in \mathbb{F}_q}[h(x) - y$ *and* $h'(x) - y'$ *have a common factor*$] = o(q^{-1})$*. Then* $F(\rho_h, \rho_{h'}) = o(1)$*.*

Proof. By Lemma 4.3.3 of [16], provided h and h' do not share a common factor, $|L_{h,0} \cap L_{h',0}| \leq q^{d-2} \deg h \deg h' \min\{\deg h, \deg h'\}$; thus we can take $\alpha = O(q^{d-2})$ with $\beta = o(q^{-1})$. By the Schwartz-Zippel Lemma (Lemma 3.3.1 in [16]), $|L_{h,0}| \leq q^{d-1} \deg h$; thus we can take $\delta = O(q^{d-1})$. Then the result follows from Lemma 8. □

Corollary 11. *Let $d \geq 2$, and suppose* $\Pr_{y \in \mathbb{F}_q}[h(x) - y$ *not absolutely irreducible*$] = o(1)$*. Then for all h' with* $\deg h' \leq \deg h$ *(other than multiples of h),* $F(\rho_h, \rho_{h'}) = o(1)$*.*

Proof. Since h is irreducible, it cannot share a common factor with h', so we can take $\alpha = O(q^{d-2})$ with $\beta = o(1)$. By Lemma 5.5.1 of [16], provided h is absolutely irreducible, $|L_{h,0}| = q^{d-1}[1 + O(q^{-1/2})]$, so we can take $\gamma = \Omega(q^{d-1})$ and $\delta = O(q^{d-1})$. Then the result follows from Lemma 9. □

Finally, we show that almost all polynomials satisfy the conditions of Corollary 11, which implies that the query complexity of typical hidden polynomial problems is $\operatorname{poly}(\log q)$.

Theorem 12. *Fix $d \geq 2$ and $t \geq 1$. Then for a fraction $1 - o(1)$ of the polynomials h in $\mathbb{F}_q[x_1, \ldots, x_d]$ of total degree t, for all h' with* $\deg h' \leq t$ *(other than multiples of h),* $F(\rho_h, \rho_{h'}) = o(1)$*.*

Proof. We show that the fraction of polynomials that are not absolutely irreducible is $O(1/q)$. Then the theorem follows by application of Corollary 11 and Markov's inequality.

The main idea is to count nontrivial factorizations of h. Let $\mathbb{F}_q(d;t)$ denote the set of d-variate polynomials over $\mathbb{F}_q$ of total degree t. If $t = 1$ then we know the states are distinguishable (since they are abelian hidden subgroup states), so we can assume $t \geq 2$. It is convenient to discuss $\mathbb{F}_q$-projectivized polynomials, i.e., equivalence classes with respect to multiplication by nonzero elements in $\mathbb{F}_q$; denote these by $\mathbb{PF}_q(d;t)$.

The number of $\mathbb{F}_q$-degrees of freedom of $\mathbb{PF}_q(d;t)$ (i.e., the number of elements of $\mathbb{F}_q$ required to specify a member of $\mathbb{PF}_q(d;t)$) is $\binom{d+t}{d} - 1$. The number of $\mathbb{F}_q$-degrees of freedom of $\mathbb{PF}_{q^k}(d;t)$ (the set of $\mathbb{F}_{q^k}$-projectivized polynomials with coefficients in $\mathbb{F}_{q^k}$) is $k\big(\binom{d+t}{d} - 1\big)$. Now we rely on the following fact: Let $h \in \mathbb{PF}_q(d;t)$. Then there is a (unique) factorization $h = h_1 \cdots h_\ell$ (for some $\ell \geq 1$) with each $h_i \in \mathbb{PF}_q(d;t)$, and of the following special form: In the (unique) factorization $h_i = \eta_{i,1} \cdots \eta_{i,k_i}$ of h_i over the algebraic closure of $\mathbb{F}_q$, for every j, the smallest field containing the coefficients of $\eta_{i,j}$ is $\mathbb{F}_{q^{k_i}}$, and the Frobenius automorphism $c \mapsto c^q$, acting on coefficients, cyclically permutes the set $\{\eta_{i,1}, \ldots, \eta_{i,k_i}\}$. (In particular, for any fixed i, the $\eta_{i,j}$ are all distinct.) Most importantly, $\eta_{i,1}$ determines all the $\eta_{i,j}$, so the number of $\mathbb{F}_q$-degrees of freedom of h_i (of degree t_i) is $k_i\big(\binom{d+t_i/k_i}{d} - 1\big)$.

There are at most t possible values for ℓ; we bound the number of factorizations by treating $\ell > 1$ and $\ell = 1$ separately.

The number of $\mathbb{F}_q$-degrees of freedom for the factorizations of h with $\ell > 1$ is upper bounded by $\max_{1 \leq t' < t} \big[\binom{d+t'}{d} + \binom{d+t-t'}{d} - 2\big]$. It suffices to show that this is $\leq \binom{d+t}{d} - 2$, hence strictly less than $\binom{d+t}{d} - 1$, the number of $\mathbb{F}_q$-degrees of freedom of $\mathbb{PF}_q(d;t)$. Fix an ordered set of size $d + t$. It has $\binom{d+t}{d}$ subsets of size d, $\binom{d+t'}{d}$ subsets of size d which avoid the last $t - t'$ elements, and $\binom{d+t-t'}{d}$ subsets of size d which avoid the first t' elements. The latter two collections have just one common element, so we need only note that for $t \geq 2$, there is at least one subset of size d which is in neither collection.

The number of $\mathbb{F}_q$-degrees of freedom for the factorizations with $\ell = 1$ is, by the earlier discussion, $\max_{k>1} \left[k\left(\binom{d+t/k}{d} - 1\right)\right]$. We again need to show that this is $\leq \binom{d+t}{d} - 2$. Fix a set of size $d+t$, and partition it into B_0 of size d and $B_1, \ldots, B_k$ each of size t/k. For any $1 \leq i \leq k$, the quantity $\binom{d+t/k}{d} - 1$ counts the subsets of size d which are contained in $B_0 \cup B_i$ but which are not equal to B_0. These are disjoint subsets. None of them includes B_0; and because $t \geq 2$, they also miss at least one other subset of size d, which intersects more than one of $B_1, \ldots, B_k$. Hence the desired inequality follows. □

Acknowledgments

We thank Dan Abramovich for his insights on the typicality of absolute irreducibility, and John Watrous for suggesting the problem of finding an efficient quantum procedure for moving amplitude from spheres to their centers. AMC received support from NSF Grant PHY-0456720 and ARO Grant W911NF-05-1-0294. LJS received support from NSF Grant PHY-0456720, ARO Grant W911NF-05-1-0294, and NSF Grant CCF-0524828. UVV received support from NSF Grant CCF-0524837 and ARO Grant W911NF-07-1-0030.

References

[1] A. Adolphson, *On the distribution of angles of Kloosterman sums*, J. Reine Angew. Math. **395** (1989), 214–220.

[2] M. Aschbacher, *Finite Group Theory*, 2nd ed., Cambridge University Press, 2000.

[3] H. Barnum and E. Knill, *Reversing quantum dynamics with near-optimal quantum and classical fidelity*, J. Math. Phys. **43** (2002), no. 5, 2097–2106.

[4] L. Carlitz, *Weighted quadratic residues over a finite field*, Can. J. Math. **5** (1953), 317–323.

[5] M. Ettinger, P. Høyer, and E. Knill, *Hidden subgroup states are almost orthogonal*, quant-ph/9901034.

[6] M. Ettinger and P. Høyer, *On quantum algorithms for noncommutative hidden subgroups*, Adv. in Appl. Math. **25** (2000), 239–251.

[7] M. Grigni, L. Schulman, M. Vazirani, and U. Vazirani, *Quantum mechanical algorithms for the nonabelian hidden subgroup problem*, Proc. 33rd STOC, 2001, pp. 68–74.

[8] S. Hallgren, C. Moore, M. Rötteler, A. Russell, and P. Sen, *Limitations of quantum coset states for graph isomorphism*, Proc. 38th STOC, 2006, pp. 604–617.

[9] A. W. Harrow and A. Winter, *How many copies are needed for state discrimination?*, quant-ph/0606131.

[10] N. M. Katz, *Gauss sums, Kloosterman sums, and Monodromy Groups*, Annals of Mathematics Studies, vol. 116, Princeton University Press, 1988. Chapter 13.

[11] A. Yu. Kitaev, *Quantum computations: Algorithms and error correction*, Russian Math. Surveys **52** (1997), no. 6, 1191–1249.

[12] A. Medrano, P. Myers, H. M. Stark, and A. Terras, *Finite analogues of Euclidean space*, J. Comput. Appl. Math. **68** (1996), 221–238.

[13] C. Moore, A. Russell, and L. J. Schulman, *The symmetric group defies strong Fourier sampling*, Proc. 46th FOCS, 2005, pp. 479–488.

[14] H. Niederreiter, *The distribution of values of Kloosterman sums*, Arch. Math. **56** (1991), 270–277.

[15] H. Salié, *Über die Kloostermanschen Summen $S(u, v; q)$*, Math. Z. **34** (1932), no. 1, 91–109.

[16] W. M. Schmidt, *Equations Over Finite Fields: An Elementary Approach*, 2nd ed., Kendrick Press, 2004.

[17] P. W. Shor, *Algorithms for quantum computation: Discrete logarithms and factoring*, SIAM J. Comput. **26** (1997), no. 5, 1484–1509.

[18] D. Simon, *On the power of quantum computation*, Proc. 35th FOCS, 1994, pp. 116–123.

[19] A. Weil, *On some exponential sums*, Proc. Natl. Acad. Sci. **34** (1948), 204–207.

A General formulation of shifted subset problems

In Section 2, we explained one way to formulate the hidden radius problem as a black box problem. In this appendix, we give an alternative definition that applies to general shifted subset problems. (It is also possible to give a general definition along the lines of Section 2, but such a definition requires certain intersection properties not required here.)

An instance of a shifted subset problem over $X := \mathbb{F}_q^d$ is specified by a subset of points S and a set of shifts T. The problem is to determine some property of S or T (or both) using a black box that hides the shifted subsets $S + t$ for $t \in T$.

To obfuscate the meanings of the shifts, we introduce a bijection $\tau : T \to T$. Furthermore, to obfuscate the meanings of the points in the subsets, we introduce a bijection $\sigma_t : S \to S$ for each $t \in T$. The black-box function $\pi : S \times T \to X$ defined as $\pi(s,t) := \tau(t) + \sigma_t(s)$ turns an input (s,t), representing an encryption of a point in the space associated with a particular shifted subset, into an explicit point $x \in X$. We associate each encrypted shift $t \in T$ with a black-box function value $f(t)$, where $f : T \to Y$ is an injection into an arbitrary finite set Y. Finally, to allow erasing the encrypted inputs (s,t), we introduce the function $g : X \times Y \to (S \times T) \cup \{\varnothing\}$ defined as

$$g(x,y) := \begin{cases} (s,t) & \exists s \in S, t \in T : \pi(s,t) = x \text{ and } f(t) = y \\ \varnothing & \text{otherwise}\,. \end{cases} \tag{29}$$

The oracle allows us to compute π, f, or g as desired.

Just as in Theorem 1, we have

Theorem 13. *Any classical computation with access to π, f, and g requires an expected exponential number of queries to obtain a $1/\operatorname{poly}(d \log q)$ bias for any single bit of information about S or T.*

The proof proceeds along the same lines as before.

However, on a quantum computer, we can prepare quantum states that encode S and T. We begin with a uniform superposition over the encrypted inputs (s,t), compute the point $x = \pi(s,t)$, compute $f(t)$, uncompute the original inputs, and finally discard the function value. This procedure results in the state (up to normalization)

$$\sum_{s\in S, t\in T} |s,t\rangle \mapsto \sum_{s\in S,\, t\in T} |s,t,\pi(s,t)\rangle \tag{30}$$

$$\mapsto \sum_{s\in S,\, t\in T} |s,t,\pi(s,t),f(t)\rangle \tag{31}$$

$$\mapsto \sum_{s\in S,\, t\in T} |\pi(s,t),f(t)\rangle \tag{32}$$

$$\mapsto |S+t\rangle \text{ where } t \text{ is uniformly random in } T\,. \tag{33}$$

In other words, we have prepared the shifted subset state

$$\rho_{S,T} := \frac{1}{|T|}\sum_{t\in T} |S+t\rangle\langle S+t|\,. \tag{34}$$

Note that we may allow the possible sets S to have different sizes, and similarly for the possible sets T. For example, we see from (5) that spheres of nonzero radius have two different sizes in odd dimensions. In such cases the black box functions can be expanded to include a symbol $\varnothing$ that is returned if the input is invalid. The above procedure can still be used provided the probability that the measurement returns the outcome $\varnothing$ is small.

B Query complexity of the HRP in odd dimensions

Proof of Theorem 3. The distribution of k is given by

$$\Pr(k|r) = \frac{1}{q|\mathcal{S}_r|}\begin{cases} |\mathcal{S}_r|^2/q^{d-1} & k = 0 \\ 1 & r\Delta(k) = 0, r \neq 0 \text{ or } \Delta(k) \neq 0 \text{ with } k \neq 0 \\ 4\cos^2(2\pi \operatorname{tr} \sqrt{r\Delta(k)}/p) & \chi(r\Delta(k)) = 1 \\ 0 & \chi(r\Delta(k)) = -1 \text{ or } r = \Delta(k) = 0 \text{ with } k \neq 0\,. \end{cases} \tag{35}$$

Now consider a pair of distinct radii r, r'. We have already described an efficient algorithm to determine $\chi(r)$ (and in particular, to decide whether $\chi(r) = 0$), so we can assume $r, r' \neq 0$. If $\chi(r) \neq \chi(r')$, then the distributions they induce have nearly disjoint support, and their total variation distance is $1 - o(1)$. Otherwise, we can rescale the spheres and the measured values of $\Delta(k)$ so that we are effectively distinguishing radius 1 from some arbitrary radius $r \neq 1$ with $\chi(r) = 1$. The minimum total variation distance between the resulting distributions is

$$\begin{aligned}
&\min_{\substack{r\in\mathbb{F}_q\setminus\{1\}\\ \chi(r)=1}} \frac{2}{q} \sum_{\substack{s\in\mathbb{F}_q\\ \chi(s)=1}} \left|\cos^2 \frac{2\pi \operatorname{tr}\sqrt{s}}{p} - \cos^2 \frac{2\pi \operatorname{tr}\sqrt{rs}}{p}\right| \\
&= \min_{r\in\mathbb{F}_q^\times\setminus\{\pm 1\}} \frac{2}{q} \sum_{s\in\mathbb{F}_q} \left|\cos^2 \frac{2\pi \operatorname{tr} s}{p} - \cos^2 \frac{2\pi \operatorname{tr} rs}{p}\right| &(36)\\
&= \min_{r\in\mathbb{F}_q^\times\setminus\{\pm 1\}} \frac{2}{q} \sum_{s\in\mathbb{F}_q} \frac{1}{2}\left|\cos \frac{4\pi \operatorname{tr} s}{p} - \cos \frac{4\pi \operatorname{tr} rs}{p}\right| &(37)\\
&\geq \min_{r\in\mathbb{F}_q^\times\setminus\{\pm 1\}} \frac{2}{q} \sum_{s\in\mathbb{F}_q} \frac{1}{4}\left|\cos \frac{4\pi \operatorname{tr} s}{p} - \cos \frac{4\pi \operatorname{tr} rs}{p}\right|^2 &(38)\\
&= \min_{r\in\mathbb{F}_q^\times\setminus\{\pm 1\}} \frac{1}{q} \sum_{s\in\mathbb{F}_q} \left(\cos^2 \frac{4\pi \operatorname{tr} s}{p} - \cos \frac{4\pi \operatorname{tr} s}{p} \cos \frac{4\pi \operatorname{tr} rs}{p}\right) &(39)\\
&= \frac{1}{q} \sum_{s\in\mathbb{F}_q} \cos^2 \frac{4\pi \operatorname{tr} s}{p} &(40)\\
&= \frac{1}{2}\,. &(41)
\end{aligned}$$

Since an arbitrary pair of radii are statistically distinguishable with constant total variation distance, poly$(\log q)$ samples are information-theoretically sufficient to identify an arbitrary radius. □

Note that the distribution (35) in the case $\chi(r\Delta(k)) = 1$ resembles the distribution induced by a well-known single-register measurement for the dihedral hidden subgroup problem [6], which has resisted attempts at efficient postprocessing.